Harvard Business School
Online

A Beginner's Guide to Data & Analytics

Contents

Data is ubiquitous. It's collected at every purchase made, flight taken, ad clicked, and social media post liked—which means it's never been more accessible to organizations.

Yet, access to data isn't all it takes to set a business on the path to success; it also takes employees who understand and know how to leverage data. There's now an increased demand for data-literate business professionals who can handle, analyze, and interpret data to drive decision-making.

"In this world of big data, basic **data literacy**—the ability to analyze, interpret, and even question data—is an increasingly valuable skill," says Harvard Business School Professor Janice Hammond in the HBS Online course Business Analytics.

With the right skills, data can allow you to gain and act on customer insights, predict future financial and market trends, and enact systemic change for social good.

Through this e-book, you'll gain an introduction to data literacy that can put you on track to be a data-driven professional. Entering into the data space as a beginner may seem daunting, but with foundational knowledge, you can build your data literacy and leverage the power of data for organizational success.

Data Science vs. Data Analytics: What's the Difference?

If you're new to the world of data, two terms you're likely to encounter are "data science" and "data analytics." While these terms are related, they refer to different things. Here's an overview of what each term means and how it applies to business.

Data Science vs. Data Analytics

Data science is the process of building, cleaning, and structuring datasets to analyze and extract meaning. **Data analytics**, on the other hand, refers to the process and practice of analyzing data to answer questions, extract insights, and identify trends.

You can think of data science as a precursor to data analysis. If your dataset isn't structured, cleaned, and wrangled, how will you be able to draw accurate, insightful conclusions? Below is a deeper dive into each field's role in business.

🎓 FACULTY INSIGHT

Putting Data Analysis to Work

HBS Professor Janice Hammond, who teaches the online course Business Analytics, implores professionals to harness the power of data analytics in a previous interview.

"Every time you do an analysis, you don't just say, 'Oh, the answer is 17. I'm done,'" she says. "You need to ask, 'What can I learn from the results of this analysis about the underlying context, about competition, about customers, about suppliers?' Managers should ask things like, 'How do the results of this analysis validate or reinforce hypotheses I had before I did the analysis?' It is equally important to ask, 'What did I learn that negates or calls into question the assumptions that I made going into the analysis?' Every analysis should be a feedback loop that deepens your learning."

Data Science in Business

In business, data science is used to collect, organize, and maintain data—often to write algorithms that make large-scale analysis possible. When designed correctly and tested thoroughly, algorithms can catch information or trends that humans miss. They can also significantly speed up the processes of gathering and analyzing data. You can use data science to:

- **Gain customer insights:** Data about your customers can reveal details about their habits, demographics, preferences, and aspirations. A foundational understanding of data science can help you make sense of and leverage it to improve user experiences and inform retargeting efforts.

- **Increase security:** You can also use data science to increase your business's security and protect sensitive information. For example, machine-learning algorithms can detect bank fraud faster and with greater accuracy than humans, simply because of the sheer volume of data generated every day.

- **Inform internal finances:** Your organization's financial team can utilize data science to create reports, generate forecasts, and analyze financial trends. Data on a company's cash flows, assets, and debts is constantly gathered, which financial analysts use to manually or algorithmically detect trends in financial growth or decline.

- **Streamline manufacturing:** Manufacturing machines gather data from production processes at high volumes. In cases where the volume of data collected is too high for a human to manually analyze it, an algorithm can be written to clean, sort, and interpret it quickly and accurately to gather insights that drive cost-saving improvements.

- **Predict future market trends:** Collecting and analyzing data on a larger scale can enable you to identify emerging trends in your market. By staying up to date on the behaviors of your target market, you can make business decisions that allow you to get ahead of the curve.

"This course was impactful, especially using case studies of real-life situations to solve complex and confusing problems. The results of this will help improve my managerial decisions within and outside organizations to minimize risks and increase profits."

Bamidele Ajisogun
Harvard Online Data Science for Business Participant

Data Analytics in Business

The main goal of business analytics is to extract meaningful insights from data that an organization can use to inform its strategy and, ultimately, reach its objectives. Business analytics can be used for:

- **Budgeting and forecasting:** By assessing a company's historical revenue, sales, and costs data alongside its goals for future growth, an analyst can identify the budget and investments required to make those goals a reality.
- **Risk management:** By understanding the likelihood of certain business risks occurring—and their associated expenses—an analyst can make cost-effective recommendations to help mitigate them.
- **Marketing and sales:** By understanding key metrics, such as lead-to-customer conversion rate, a marketing analyst can identify the number of leads their efforts must generate to fill the sales pipeline.
- **Product development (or research and development):** By understanding how customers reacted to product features in the past, an analyst can help guide product development, design, and user experience in the future.

Now that you have an understanding of the difference between data science and data analytics and their business applications, you can dive deeper into the world of data. The next section covers basic concepts to build your data literacy.

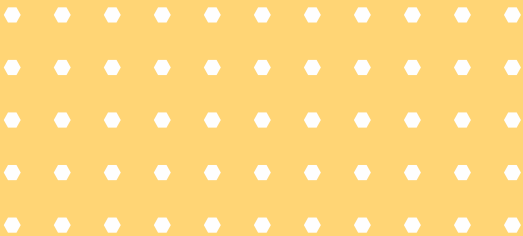## BUSINESS INSIGHT

4 Types of Analytics

Analytics is used to extract meaningful insights from data that can drive decision-making and strategy formulation. There are four types of analytics you can leverage depending on the data you have and the type of knowledge you'd like to gain.

1. **Descriptive analytics** looks at data to examine, understand, and describe something that's already happened.

2. **Diagnostic analytics** goes deeper than descriptive analytics by seeking to understand the "why" behind what happened.

3. **Predictive analytics** relies on historical data, past trends, and assumptions to answer questions about what will happen in the future.

4. **Prescriptive analytics** identifies specific actions an individual or organization should take to reach future targets or goals.

Data Literacy 101: Familiarizing Yourself with the Data Landscape

Data literacy *is the ability to read, understand, and utilize data in different ways. Being data-literate can help non-data professionals read and understand data, and use it to inform their decision-making. Here's a breakdown of key areas of the data landscape.*

Data Ecosystem & Lifecycle

The term **data ecosystem** refers to the programming languages, packages, algorithms, cloud-computing services, and general infrastructure an organization uses to collect, store, analyze, and leverage data. No two organizations leverage the same data in the same way. As such, each organization has a unique data ecosystem.

While the data ecosystem encompasses everything that handles, organizes, and processes data, the **data life cycle** describes the path data takes from when it's first generated to when it's interpreted into actionable insights. This life cycle can be split into eight steps: generation, collection, processing, storage, management, analysis, visualization, and interpretation.

A data project's steps are often described as a cycle because the lessons learned and insights gleaned from one project typically inform the next. In this way, the final step of the process feeds back into the first, enabling you to start again with new goals and learnings.

8. INTERPRETATION
1. GENERATION
2. COLLECTION
3. PROCESSING
4. STORAGE
5. MANAGEMENT
6. ANALYSIS
7. VISUALIZATION
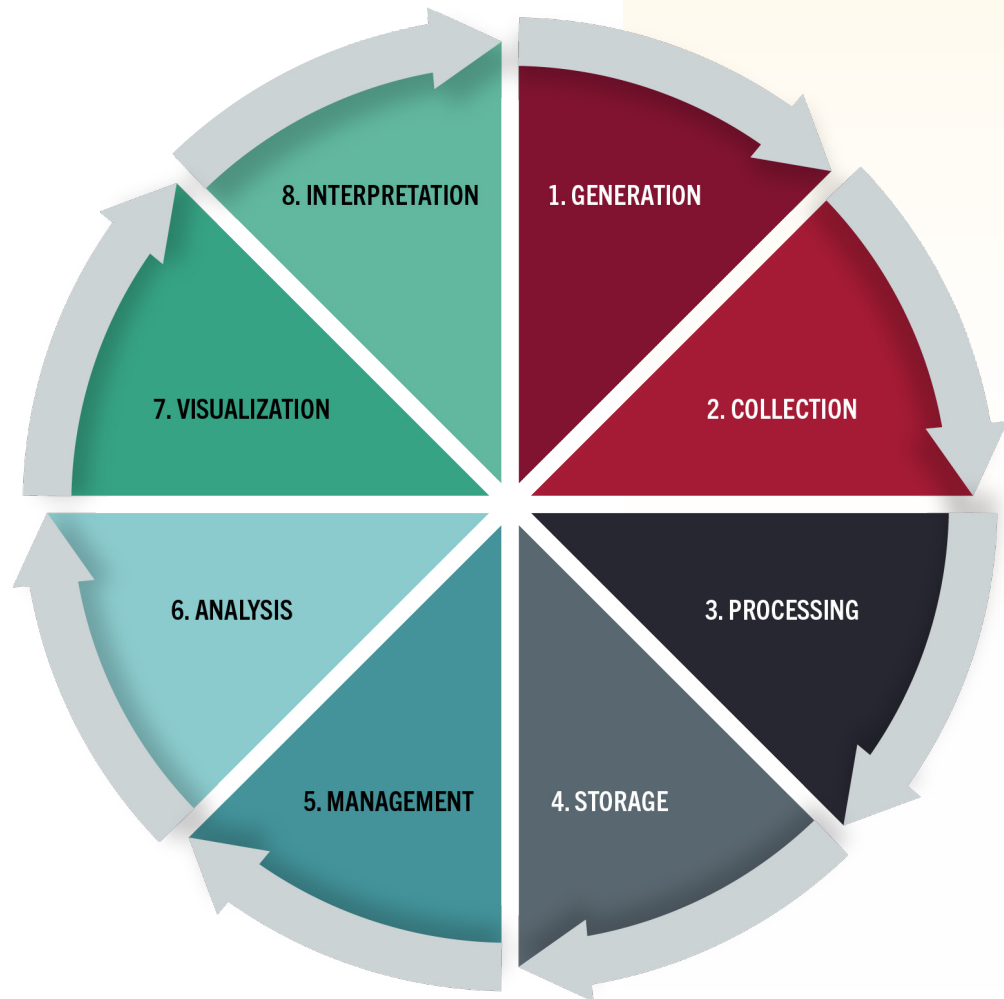
Data Privacy & Ethics

Data privacy, also known as information privacy, is a subcategory of data protection that encompasses the ethical and legal obligation to protect access to **personally identifiable information (PII)**, which is any information that can be linked to a specific individual. Some examples of PII include full name, address, Social Security number, and passport number.

In the Harvard Online course Data Science Principles, taught by Professor Dustin Tingley, it's explained that data privacy is made up of three key questions:

1. *What data is collected?*

2. *How is the data stored?*

3. *Who can access the data?*

The ethics of data privacy can be boiled down to the fact that an individual's consent is necessary to collect, store, and use their personal information.

As a data handler, you have a responsibility to be transparent with your subjects about your intentions, what their data will be used for, and who will have access to it. In addition, you need to ensure your use of data doesn't cause harm to an individual or group of people. This is referred to as **disparate impact** and is unlawful.

It's imperative that customer privacy is at the forefront of your mind during the data analysis process. Data scientists and business professionals alike need to maintain a set of ethical tenets that guide their data privacy policies. Two pointers: Always ask for consent before collecting data, and research the privacy laws that apply to your geographic regions and market segments.

While the privacy and ethical use of data are everyday efforts, knowing that your subjects' safety and rights are intact is worth the work.

Data Literacy 101: Familiarizing Yourself with the Data Landscape

Data Integrity

Data integrity is the accuracy, completeness, and quality of data as it's maintained over time and across formats. Preserving the integrity of your company's data is a constant process.

Threats to a dataset's integrity include:

- **Human error**: For instance, accidentally deleting a row of data in a spreadsheet.

- **Inconsistencies across formats**: For instance, a dataset in Microsoft Excel that relies on cell referencing may not be accurate in a different format that doesn't allow those cells to be referenced.

- **Collection error**: For instance, data collected is inaccurate or lacking information, creating an incomplete picture of the subject.

- **Cybersecurity or internal privacy breaches**: For instance, someone hacks into your company's database with the intent to damage or steal information, or an internal employee damages data with malicious intent.

To maintain your datasets' integrity, diligently check for errors in the collection, formatting, and analysis phases, monitor for potential breaches, and educate your team about the importance of data integrity.

"This is a topic that people in any industry should have at least basic knowledge of in order to create more efficient and competitive businesses, tools, and resources."

Carlos E. Sapene
Harvard Online Data Science Principles Participant

Understanding the way data must be handled, processed, and protected is an imperative step on your way to becoming data-literate. The next section explores the data skills you need to succeed and how to build them.

Building Your Data & Analytical Skill Set

Data and analytics are important skills to have in any industry because they enable you to support decisions with data, learn more about customers, and predict future trends. Here are seven skills for success and ways to develop them.

7 Data & Analytics Skills You Need

1. Critical Thinking

If you're interested in using data to solve business problems, you need to be adept at thinking critically about challenges and solutions. While data can provide many answers, it's nothing without a human's discerning eye.

"From the first steps of determining the quality of a data source to determining the success of an algorithm, critical thinking is at the heart of every decision data scientists—and those who work with them—make," Tingley says in the Harvard Online course Data Science Principles. "Data science is a discipline that's built on a foundation of critical thinking."

2. Hypothesis Formation and Testing

At the heart of data and analytics is the desire to answer questions. The proposed explanations for these leading questions are called **hypotheses**, which must be formed before analysis takes place.

An example of a hypothesis is, "I predict that a person's likelihood of recommending our product is directly proportional to their reported satisfaction with the product." You predict the data will show this trend and must prove or disprove the hypothesis through analysis. Without a hypothesis, your analysis has no clear direction.

3. Data Wrangling

Data wrangling is the process of cleaning raw data in preparation for analysis. It involves identifying and resolving mistakes, filling in missing data, and organizing and transferring it into an easily understandable format.

This is an important skill for anyone dealing with data to acquire because it leads to a more efficient and organized data analysis process. You can extract valuable insights from data more quickly when it's cleaned and in its optimal viewing format.

4. Mathematical Ability

You don't have to be a mathematician to become data literate, but strong math skills become increasingly important as you deal with more complex analyses.

A seasoned data professional needs a solid understanding of statistics, probability, linear algebra, and multivariable calculus. Data scientists often call on statistical methods to find structure in data and make predictions, and linear algebra and calculus can make machine-learning algorithms easier to comprehend.

If you're not a data scientist or analyst, your work may not require you to understand the more complex mathematical concepts, but having a basic understanding of statistics can go a long way.

5. Data Visualization

It's crucial to know how to transform raw data into compelling visuals that tell a story.
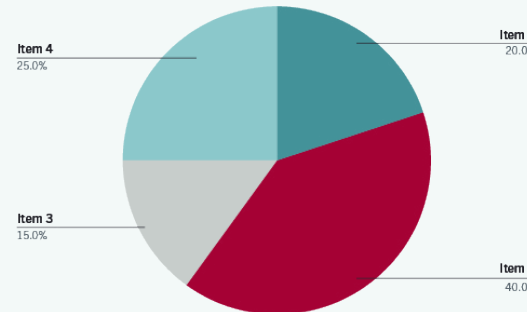
Rather than simply presenting a list of values to your stakeholders, it's more effective to visually communicate data in a way that's easily digestible. Some popular data visualization techniques that all business professionals should know include pie charts, bar charts, and histograms.

To create these visualizations, use a **data visualization tool**, a form of software designed to present data. Each tool's capabilities vary but, at their most basic, allow you to input a dataset and visually manipulate it. Most, but not all, come with built-in templates you can use to generate basic visualizations. Examples include Microsoft Excel and Power BI, Google Charts, Tableau, Zoho Analytics, Data Wrapper, and Infogram.
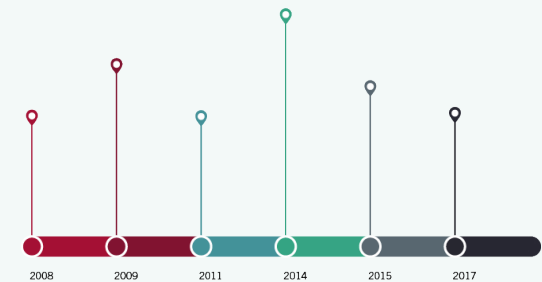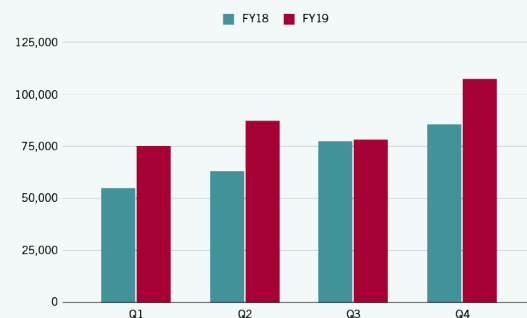
DATA VISUALIZATION TECHNIQUES

Pie Chart

Item 4
25.0%

Item 1
20.0%

Item 3
15.0%

Item 2
40.0%

Timeline

2008 2009 2011 2014 2015 2017

Bar Graph

FY18 FY19

125,000
100,000
75,000
50,000
25,000
0
Q1 Q2 Q3 Q4

Area Graph

2017 2018 2019

50,000
40,000
30,000
20,000
10,000
0
January February March April May June July
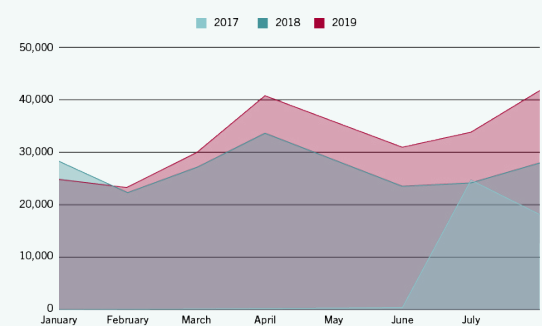
6. Programming

Programming languages, like Python and R, are commonly used to solve complex statistical problems with data. Proficiency in a database querying language, like SQL, can also help you more easily extract and change data in a database.

While programming skills are immensely valuable, they're not necessary for beginners dabbling in data. It's more important to focus on effectively analyzing and visualizing data to draw conclusions.

7. Machine Learning

As artificial intelligence grows in popularity, machine learning is a highly valuable skill for professionals working with big data.

Machine learning refers to the use of computer algorithms that automatically learn from and adapt in response to data. Some business applications of machine learning include risk management, performance analysis, trading, and automation.

Even if you're not responsible for writing code, knowing the basics of machine learning can help you gain a deeper understanding of your organization and boost efficiency through automation.

How to Improve Your Skills

1. Embrace the Challenge

The first step is to confront any mental barriers surrounding your ability to learn and develop data skills.

"Data science isn't scary, and it shouldn't be intimidating," says Professor Yael Grushka-Cockayne in the Harvard Online course Data Science for Business. "Combining your business know-how and intuition with data science can set you and your company up for success."

Although data science has a reputation for being code-based and complex, its concepts are accessible if you have the desire and drive to learn and put in the work.

"Some folks are under the impression that unless they were trained as a data scientist and have been coding for years, they don't stand a chance," Grushka-Cockayne says in a recorded webinar. "But that's not accurate. It's never too late."

2. Consider Opposing Viewpoints

While engaging with opposing viewpoints can help you expand your perspective, combat bias, and show your fellow employees their opinions are valued, it can also be a useful way to practice analytical skills.

When analyzing data, it's crucial to consider all possible interpretations and avoid getting stuck in one way of thinking.

For instance, imagine you track users who click a button on your site to download an e-book. The data shows that the user's age is positively correlated with their likelihood to click the button; as age increases, downloads increase. At first glance, you may interpret this trend to mean that a user downloads the e-book because of their age.

This conclusion doesn't take into consideration the variables that change with age. For instance, perhaps the real reason older users are more likely to download the e-book is their higher level of responsibility at work, average income, or likelihood of being parents.

This example illustrates the need to consider multiple interpretations of data, and it specifically shows the difference between **correlation** (the trending of two or more variables in the same direction) and **causation** (when a trend in one variable causes a trend to occur in one or more other variables).

To practice this skill, question your assumptions and ask others for their opinions. The more you actively engage with different viewpoints, the less likely you are to get stuck in a one-track mindset when analyzing data.

3. Play Games or Brain Teasers

If you're looking to sharpen your skills on a daily basis, there are many simple, enjoyable ways to do so.

Games, puzzles, and stories that require visualizing relationships between variables, examining situations from multiple angles, and drawing conclusions from known data points can help build the skills needed to analyze data. Some fun ways to practice analytical thinking include crossword puzzles, riddles, mystery novels, Sudoku, and logic puzzles.

These options can supplement your analytics coursework and on-the-job experience. Some also allow you to spend time with friends or family. Try engaging with one each day to hone your analytical mindset.

4. Learn From Real-World Examples

By exploring how other business professionals use data to solve problems, you can imagine what you'd do in their scenarios, evaluate the impact of their actions, and put that knowledge into practice.

Case-based learning is one of the foundations of HBS Online courses. Each offers several examples of business leaders at well-known organizations who used data to solve problems.

"You need to make it real," Grushka-Cockayne says in a recorded webinar. "You need to make it relevant and ask, 'Why do I care about this?' or 'Why do I want to look at a summary statistic?' or 'How is this going to be meaningful for a specific decision?' By exposing yourself to cases from various industries, the richness of the variety allows you to put yourself in a decision-maker's position and understand how actual decisions get made."

"I enjoyed how the principles in the course were grounded in concrete applications to the real world. I majored in engineering and minored in economics as an undergraduate, but the content showed me new ways for using the theoretical foundations I already had."

Carolina Ragolta
HBS Online Business Analytics Participant

5. Find a Community

On your path to learning data and analytics, a community of fellow professionals with the same goal can be a motivating, supportive force. You can turn to online forums, social media, affinity groups within your organization or geographic area, or a cohort of learners in an online class.

For instance, the HBS Online Community is a global network of learners dedicated to furthering their business education. Having this kind of community allows you to ask for feedback and advice, and engage with others on new concepts as you work toward goals.

6. Engage with and Ask Big Questions of Your Data

To engage with data further, ask questions. With each question comes an opportunity to uncover more insight and gain skills. These questions can lead you to learn a new coding language, analysis method, regression, or visualization tool to help solve a specific business problem. Here are examples of questions to ask yourself when handling data:

- *What am I hoping to understand?*
- *What do I need to know to make a certain business decision?*
- *What story is this data telling?*
- *What do the relationships between variables mean for _____ at my company?*
- *What if _____ changed? Which variables, trends, or forecasts would be impacted?*
- *What needs to change in the data to get the desired outcome?*
- *Why does the data trend in this direction, and what does it mean for the future?*
- *How can I further analyze the data to get the answers needed to make important decisions?*

Make the data work for you by determining what you need to know and the best way to answer questions using available data. Improving your data and analytical skills is an ongoing process, and each experience offers a chance to learn more.

⚲ BUSINESS INSIGHT

The Data-Driven Decision-Making Framework

There are various tools and frameworks that can be beneficial when using data. One is the data-driven decision-making framework, which is explained in the Harvard Online course Data Science for Business. This framework presents six steps for using data to drive business decisions:

Data-Driven Decision-Making Framework
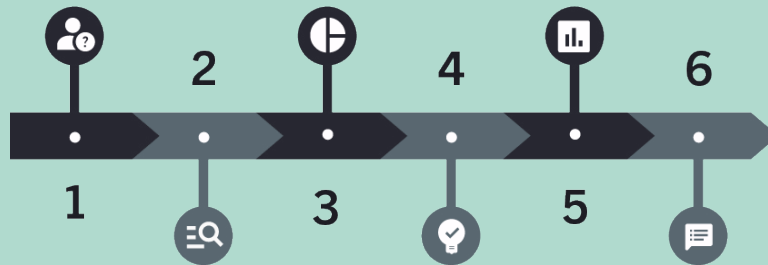
1. **Understand the business problem:** What are you looking to understand or accomplish?

2. **Wrangle data:** Clean, validate, and organize the data.

3. **Create visualizations:** Present the data in a way that shows trends and relationships of interest.

4. **Generate hypotheses:** Formulate predictions based on emerging trends.

5. **Conduct analysis:** Run statistical tests to determine if your hypotheses are correct.

6. **Communicate results:** Present your findings in the context of the original business problem.

Understanding data frameworks like this one gives you the capability to take a raw dataset, interpret its story, and use it to answer relevant business questions.

Like all skills, those pertaining to data and analytics require practice to master. If you're ready to take your learning to the next level, consider enrolling in a data science or data analytics course. The next section outlines some options.

Which Data & Analytics Course Is Right for You?

Whether you want to learn the basics, brush up on your skills, or expand your knowledge, taking a data or analytics course is an effective way to improve. Take stock of your current data and analytics knowledge and the goals you have for the future. Then, explore the course comparison table to help determine which HBS Online or Harvard Online Data and Analytics course is right for you.

Data & Analytics Courses

Basic Info	What You'll Learn	Who Will Benefit
Business Analytics Harvard Business School Online **Faculty:** Jan Hammond **Course Length:** 8 Weeks Learn More	• Interpret data and analyze relationships between variables • Develop and test hypotheses and draw conclusions • Implement regression analysis and other analytical techniques in Excel	• College students and recent graduates aiming to start their career with analytical skills • Those considering graduate school for business or a data-focused discipline • Mid-career professionals who want to gain the skills to interpret data and use it to drive decision-making
Data Science Principles Harvard Online **Faculty:** Dustin Tingley **Course Length:** 4 Weeks Learn More	• Understand the modern data science landscape and technical terminology for a data-driven world • Recognize major concepts and tools in the field of data science and determine where they can be appropriately applied • Appreciate the importance of data organization, privacy, and ethics	• Students and recent graduates looking to gain a foundational understanding of the data landscape • Early- and mid-career professionals who want the skills to handle their organization's data • Marketing and product management professionals who want to leverage the power of data
Data Science for Business Harvard Online **Faculty:** Yael Grushka-Cockayne **Course Length:** 4 Weeks Learn More	• Gain a foundational understanding of data science and how it relates to organizational success • Create a data-driven framework for your organization • Understand key techniques and learn how to read basic code to comprehend the syntax that informs data requests	• Aspiring managers aiming to level up their data-driven decision-making skills • Managers and rising leaders who want to learn to increase revenue while mitigating risk • Product managers who want a better handle on customer data and its applications

Basic Info	What You'll Learn	Who Will Benefit
Big Data for Social Good Harvard Online **Faculty:** Raj Chetty **Course Length:** 4 Weeks Learn More	• Examine historical income, education and family support, and geography to understand how these economic factors lead to upward mobility • Understand how big data is used to identify the causes of socioeconomic disparities and how it can lead to evidence-based action and outcomes	• Community activists and nonprofits that want to leverage data in their efforts toward social good • Policy advocates and public sector employees seeking data to address systemic changes • Students and recent graduates looking to make an impact using data
Data Privacy and Technology Harvard Online **Faculty:** Michael D. Smith and Jim Waldo **Course Length:** 5 Weeks Learn More	• Understand the various attempts to define privacy and the ongoing conflict of privacy laws and norms within technology advancements • Explore the impacts of emerging technologies on the future of privacy, protection, and law • Think critically about privacy issues from multiple angles, exploring policy, cultural, and societal impacts	• IT, programming, and marketing professionals who handle data and need to understand privacy implications • Managers and decision-makers aiming to understand the intersection of privacy and technology • Advocates and policy makers interested in consumer privacy and technology

A Foundation in Data

After reading this e-book, you have a basic level of data literacy. You now understand the multifaceted data landscape, along with the skills necessary for data analytics and how to hone them. Finally, you took stock of your goals and determined which online course could be the next step in your data and analytics education.

Data is a powerful tool. As you move forward, you can level up your skills and use it to drive critical business decisions.

To learn more about what HBS Online can do for you, visit **online.hbs.edu**.

Harvard
Business
School
Online